ELSEVIER

Linear Algebra and its Applications 325 (2001) 75–80

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

A note on constrained convergence

Gorazd Lešnjak [1]

Faculty of Electrical Engineering and Computer Science, University of Maribor, Smetanova 17, P.O. Box 238, 2000 Maribor, Slovenia

Received 4 February 2000; accepted 22 August 2000

Submitted by F. Uhlig

Abstract

The entrywise convergence of the constrained powers PA^nQ of matrices is investigated. A new proof of the characterization of convergence, obtained recently by R.E. Hartwig and P. Šemrl [Rocky Mt. J. Math. 29 (1) (1999) 177], is presented. As an important tool we use a theorem from the theory of Diophantine approximation. © 2001 Elsevier Science Inc. All rights reserved.

AMS classification: 15A24; 15A18

Keywords: Entrywise convergence; Constrained powers of matrices; Effective spectrum

1. Introduction

The convergence of sequences of matrices of the form PA^nQ has been thoroughly investigated in [2] where its importance, for example in the Picard iteration, was described. In what follows, A will denote $N \times N$ matrix with complex entries. P and Q are given complex matrices of any appropriate sizes, say $J \times N$ and $N \times K$, respectively. Let $\mathbb{C}[t]$ stand for the algebra of all polynomials in t with complex coefficients. In [2] the *effective spectrum* of A, relative to P and Q (denoted by $\sigma_{PQ}(A)$) was introduced as the set of all zeroes of the monic polynomial Φ_{PQ}, a generator of the ideal

$$W_{PQ} = \{f \in \mathbb{C}[t] \mid Pg(A)f(A)Q = 0 \text{ for all } g \in \mathbb{C}[t]\}.$$

Email address: gorazd.lesnjak@uni-mb.si (G. Lešnjak).

[1] This work was supported by the Ministry of Science and Technology of Slovenia.

PII: S 0 0 2 4 - 3 7 9 5 (0 0) 0 0 3 2 6 - 8

This polynomial divides the minimal polynomial of A which we write in the form $\Phi(t) = \prod_{i=1}^{I}(t-\lambda_i)^{m_i}$ where, as usual, λ_i, $1 \leqslant i \leqslant I$, denote its zeroes, the distinct eigenvalues of A. Then $\Phi_{PQ}(t) = \prod_{i=1}^{I}(t-\lambda_i)^{\kappa_i}$ with $0 \leqslant \kappa_i \leqslant m_i$. Hartwig and Šemrl [2] call the zeroes λ_i of Φ_{PQ}, for which $\kappa_i \geqslant 1$, the *effective eigenvalues* and the corresponding exponents κ_i the *effective indices* of A, relative to P and Q. With these notions at hand we state the characterization of convergence of constrained powers [2, assertion (v) of Corollary 1, p. 191] in the following form.

Theorem 1. *The sequence* (PA^nQ), $n \in \mathbb{N}$, *converges if and only if for* $\lambda \in \{z \in \mathbb{C} \,|\, |z| \geqslant 1\}$ *we have*

$$\lambda \in \sigma_{PQ}(A) \iff \lambda = 1 \text{ and the corresponding effective index is } 1.$$

The proof of the only if part of the above assertion presented in [2] involves some nontrivial calculations. Our aim here is to give a different proof based on analysis of the growth of entries.

To achieve our goal we first study relations between certain ideals and subspaces arising naturally in the context of entrywise convergence. This enables us to show in the second step that for $n > N$ each entry of PA^nQ can be written as a linear combination of powers of nonzero eigenvalues of A with coefficients being polynomials in n having fixed degrees. To reach the desired conclusion about the convergence at a particular entry we finally combine some elementary analysis with a theorem from the theory of Diophantine approximation.

Let us mention that the entrywise convergence of matrix powers and some related questions have been studied also in the recent paper of Uhlig [5].

2. The structure of an entry

We start by fixing some necessary notation. Suppose p_j^*, $1 \leqslant j \leqslant J$, are the rows of P and q_k, $1 \leqslant k \leqslant K$, are the columns of Q. Then $p_j^* A^n q_k$ is the j,k-entry of PA^nQ. Similar to the general case but with *jk* as indices one can define ideals W_{jk} with generators Φ_{jk}, effective spectra σ_{jk}, effective eigenvalues and indices κ_{ijk} relative to the column and row chosen. These notions are related in the following way: $W_{PQ} = \bigcap_{jk} W_{jk}$, Φ_{PQ} is the least common multiple of polynomials Φ_{jk}, $1 \leqslant j \leqslant J$, $1 \leqslant k \leqslant K$, $\sigma_{PQ}(A) = \bigcup_{jk} \sigma_{jk}$ and $\kappa_i = \max\{\kappa_{ijk} \,|\, 1 \leqslant j \leqslant J,\ 1 \leqslant k \leqslant K\}$. Further, for each j, $1 \leqslant j \leqslant J$, we define the subspace of $X = \mathbb{C}^N$

$$V_j = \left\{ g(A)^* p_j \mid g \in \mathbb{C}[t] \right\},$$

with * standing for conjugate transpose. It follows that for each k, $1 \leqslant k \leqslant K$, we have

$$W_{jk} = \left\{ f \in \mathbb{C}[t] \mid f(A) q_k \in V_j^{\perp} \right\}.$$

Here, $\perp$ denotes the orthogonal complement with respect to the usual inner product in X. In accordance with the factorization of Φ, the minimal polynomial of A, we decompose X into A-invariant subspaces $X = \bigoplus_{i=1}^{I} X_i$ with standard convention $X_i = \ker(A - \lambda_i I)^{m_i}$. This implies that for each k we have $q_k = \sum_i q_{ki}$ with unique components $q_{ki} \in X_i$. All sums in this note have a finite number of terms and we explicitly write down the corresponding ranges of summation only when it is necessary. Now we can state:

Proposition 1. *If* $\lambda_i \in \sigma(A)$, *then*

$$\kappa_{ijk} = \min\left\{s \mid (A - \lambda_i I)^s q_{ki} \in V_j^{\perp}\right\}.$$

Proof. Suppose $q_{ki} \notin V_j^{\perp}$. The set $U_{ijk} = \{p \in \mathbb{C}[t] \mid p(A)q_{ki} \in V_j^{\perp}\}$ is an ideal because of the A-invariance of $V_j^{\perp}$. Clearly, $q_{ki} \in X_i = \ker(A - \lambda_i I)^{m_i}$ implies that $(t - \lambda_i)^{m_i}$ belongs to U_{ijk}. Let $(t - \lambda_i)^{\mu_{ijk}}$ be the monic generator of U_{ijk}. Then, $\mu_{ijk} = \min\{s | (A - \lambda_i I)^s q_{ki} \in V_j^{\perp}\}$ and $1 \leqslant \mu_{ijk} \leqslant m_i$. Furthermore, $\Phi_{jk} \in U_{ijk}$ gives us $\mu_{ijk} \leqslant \kappa_{ijk}$ and $W_{jk} \subseteq U_{ijk}$. In the case $q_{ki} \in V_j^{\perp}$ (including the possibility $q_{ki} = 0$) we define $\mu_{ijk} = 0$ (and $U_{ijk} = \mathbb{C}[t]$). On the other hand, the subspaces X_i and $V_j^{\perp}$ are invariant under A and hence, $W_{jk} = \bigcap_i U_{ijk}$ which further yields that

$$\prod_i (t - \lambda_i)^{\kappa_{ijk}} = \Phi_{jk}(t) = \operatorname{lcm}\left\{(t - \lambda_i)^{\mu_{ijk}} \mid 1 \leqslant i \leqslant I\right\} = \prod_i (t - \lambda_i)^{\mu_{ijk}}.$$

Here, we used the common abbreviation lcm for the least common multiple. $\square$

It is now clear that any components q_{ki} of q_k, which are already in $V_j^{\perp}$, do not contribute to κ_{ijk}. This explains the usage of the word *effective* for λ_i with $\kappa_{ijk} > 0$, for in this case the eigenvalue might be really important for the convergence of the sequence at the *jk*-entry.

To simplify the rest of our considerations, we introduce the following notation: $q_{ki}^{(0)} = q_{ki}$ and, recursively for $s \in \mathbb{N}$, $q_{ki}^{(s)} = (A - \lambda_i I) q_{ki}^{(s-1)} = (A - \lambda_i I)^s q_{ki}$. For $s \geqslant m_i$ this is equal to 0 by the definition of q_{ki}. Moreover, we know that for $s \geqslant \kappa_{ijk}$ we have $p_j^* q_{ki}^{(s)} = 0$ which we shall use in the proof below.

Proposition 2. *For all nonzero eigenvalues of A that are effective relative to the jk-entry there exist polynomials* p_{ijk} *of degree* $\kappa_{ijk} - 1$ *such that for all* $n > N$ *the equality*

$$p_j^* A^n q_k = \sum_{\lambda_i \neq 0} \lambda_i^n p_{ijk}(n)$$

holds true.

Proof. Using the notation introduced above, for $n \geqslant m_i$, we have by binomial expansion

$$\begin{aligned}A^n q_{ki} &= [\lambda_i I + (A-\lambda_i I)]^n q_{ki}\\ &= \lambda_i^n q_{ki}^{(0)} + \binom{n}{1}\lambda_i^{n-1} q_{ki}^{(1)} + \cdots + \binom{n}{m_i-1}\lambda_i^{n-m_i+1} q_{ki}^{(m_i-1)}.\end{aligned}$$

Multiplying by p_j^* from the left and rearranging yields

$$\begin{aligned}p_j^* A^n q_{ki} = \lambda_i^n \Bigg(& p_j^* q_{ki}^{(0)} + \binom{n}{1}\lambda_i^{-1} p_j^* q_{ki}^{(1)}\\ & + \cdots + \binom{n}{\kappa_{ijk}-1}\lambda_i^{-\kappa_{ijk}+1} p_j^* q_{ki}^{(\kappa_{ijk}-1)}\Bigg).\end{aligned}$$

The expression to the right of λ_i^n is clearly a polynomial in n, denoted $p_{ijk}(n)$, of degree $\kappa_{ijk}-1$. For $n \geqslant m_i$ any components q_{ki} of q_k associated with a zero eigenvalue vanish, hence the jkth entry becomes

$$p_j^* A^n q_k = p_j^* A^n \left(\sum_i q_{ki}\right) = \sum_{\lambda_i \neq 0} \lambda_i^n p_{ijk}(n),$$

which proves the assertion. □

3. The analysis of growth

In this section the final steps to prove Theorem 1 will be given.

From elementary analysis it is known that sequences of the form $p_m(n)a^n$, where p_m stands for a polynomial with constant complex coefficients of degree m and a denotes a complex number, converge to 0 if and only if $|a| < 1$. For $|a| \geqslant 1$ such sequences diverge if $a \neq 1$ or $m \geqslant 1$. Hence, when analyzing the convergent or divergent behavior of entries one has to give special attention to terms that contain effective eigenvalues with modulus not less than 1. It will be convenient to assume that effective eigenvalues are ordered according to descending absolute value: $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots$. We now consider several cases.

First, suppose we have an eigenvalue λ_1 such that $|\lambda_1| > |\lambda_i|$ for every $i \neq 1$. The jkth entry can be put in the form

$$p_j^* A^n q_k = \lambda_1^n p_{1jk}(n)\left(1 + \sum_{i>1} (\lambda_i/\lambda_1)^n p_{ijk}(n)/p_{1jk}(n)\right).$$

From $|\lambda_i/\lambda_1| < 1$ we see that the sum for $i > 1$ converges to 0 and thus $p_j^* A^n q_k$ behaves as $\lambda_1^n p_{1jk}(n)$. The degree of $p_{1jk}(n)$ is $\kappa_{1jk}-1$, hence $(p_j^* A^n q_k)$ converges iff $\lambda_1 = 1$ and $\kappa_{1jk} = 1$.

It remains to prove that the sequence $p_j^* A^n q_k$ diverges if we have different eigenvalues $\lambda_1, \ldots, \lambda_s$ with the same absolute value $\rho \geqslant 1$ being maximal among absolute values of the effective eigenvalues at the j, k-entry.

The first possibility that we encounter is that among these eigenvalues there is only one, say λ_1, having the greatest effective index κ_{1jk} (where κ_{ijk} is greater than 1). In the right-hand factor of

$$\lambda_1^n p_{1jk}(n)\left(1+\sum_{i>1}(\lambda_i/\lambda_1)^n p_{ijk}(n)/p_{1jk}(n)\right)$$

the first $s-1$ quotients $p_{ijk}(n)/p_{1jk}(n)$ converge to 0. Also, for $i>s$ we have $|\lambda_i/\lambda_1|<1$. Hence, this case reduces to the behavior of $\lambda_1^n p_{1jk}(n)$. But now $\kappa_{1jk}>1$, and the sequence at the jk-entry does not converge.

The remaining possibility is that besides $|\lambda_1|=\cdots=|\lambda_s|=\rho\geqslant 1$ some of these eigenvalues, say the first r of them, have also the same greatest effective index $\kappa=\kappa_{ijk}>1$, for $1\leqslant i\leqslant r$. We take out ρ^n in order to analyze the behavior of angular parts only. We denote the coefficients of the leading terms of polynomials $p_{1jk},\ldots,p_{rjk}$ by $m_1,\ldots,m_r$, and the arguments of the eigenvalues $\lambda_1,\ldots,\lambda_r$ by $2\pi a_1,\ldots,2\pi a_r$. Hence, we examine the behavior of the sequence of the form

$$\begin{aligned}\rho^n[n^{\kappa-1}((m_1\mathrm{e}^{2\pi \mathrm{i}na_1}+\cdots+m_r\mathrm{e}^{2\pi \mathrm{i}na_r})+\cdots+(\lambda_{r+1}/\rho)^n p_{r+1jk}(n)/n^{\kappa-1}+\cdots)\\ +(\lambda_{s+1}/\rho)^n(\cdots)+\cdots].\end{aligned}$$

We may and do assume all a_i's are nonnegative for $1\leqslant i\leqslant r$. It is evident that the terms in the second line converge to 0. The first line begins with the term containing the highest power of n, while all the other terms, belonging either to the first r eigenvalues or not, contain lower powers of n. Our aim now is to show that the factor $m_1\mathrm{e}^{2\pi \mathrm{i}na_1}+\cdots+m_r\mathrm{e}^{2\pi \mathrm{i}na_r}$ is bounded from 0 for an infinite sequence of integers thus yielding the divergence of the sequence $(p_j^*A^nq_k)$. The proof splits into two cases depending on whether all a_i's are rational or not.

Let a_i, $1\leqslant i\leqslant r$, be rational. Write $a_i=\alpha_i/\beta_i$ with $\beta_i\in\mathbb{N}$ and $\gcd(\alpha_i,\beta_i)=1$, where gcd stands for the greatest common divisor. If $n\in\mathbb{N},\ n>N$, is a multiple of $\beta=\operatorname{lcm}\{\beta_i \mid 1\leqslant i\leqslant r\}$, then

$$m_1\mathrm{e}^{2\pi \mathrm{i}na_1}+\cdots+m_r\mathrm{e}^{2\pi \mathrm{i}na_r}=m_1+\cdots+m_r=c_0.$$

If $c_0\neq 0$, we are done. If not, then for $n=m\beta+l$, $m\in\mathbb{N}$ and $0\leqslant l<r$ we define $c_l=m_1\mathrm{e}^{2\pi \mathrm{i}na_1}+\cdots+m_r\mathrm{e}^{2\pi \mathrm{i}na_r}=\sum_i m_i(\lambda_i/\rho)^l$. The eigenvalues λ_i, $1\leqslant i\leqslant r$, are distinct, hence such are quotients $\mu_i=\lambda_i/\rho$ and consequently, $r\leqslant\beta$. The coefficients μ_i^l in identities $c_l=\sum_i m_i\mu_i^l$, $0\leqslant l<r$, form the Vandermonde matrix which has in our case nonzero determinant. Taking into account that m_i, $1\leqslant i\leqslant r$, are nonzero this shows that at least one of c_l's is nonzero. For such l the subsequence of $(p_j^*A^nq_k)$ determined by $n=m\beta+l$, $m\in\mathbb{N}$, certainly diverges.

Next, suppose that at least one of the arguments a_i, $1\leqslant i\leqslant r$, is irrational. To finish the proof in this case, we shall use the following theorem of Dirichlet on simultaneous homogeneous approximation (see [3, p. 27, Theorem 1A and Corollary 1B], [4, p. 34, Theorem 1B], for the version below, and [1, p.12, Theorem 2], for the version with an improved estimate).

Theorem 2. *Let $a_1, \ldots, a_r$ be real numbers, at least one of them being irrational. There exist infinite sequences $(\alpha_n^{(i)}) \subseteq \mathbb{Z}$, $1 \leqslant i \leqslant r$, and $(\beta_n) \subseteq \mathbb{N}$ such that for all $n \in \mathbb{N}$ we have* $\gcd(\beta_n, \alpha_n^{(1)}, \ldots, \alpha_n^{(r)}) = 1$ *and*

$$\left| a_i - \frac{\alpha_n^{(i)}}{\beta_n} \right| < \frac{1}{\beta_n^{1+1/r}} \quad \textit{for all } i,\ 1 \leqslant i \leqslant r.$$

In other words, there exists an increasing sequence (β_n) of integers such that for all i, $1 \leqslant i \leqslant r$, we have $|\beta_n a_i - \alpha_n^{(i)}| < \beta_n^{-1/r}$. Choosing a subsequence (k_n) of (β_n), satisfying $r < k_{n+1} - k_n$, as the exponents for powers of A, one concludes that for n great enough we get

$$c_0^{(n)} = m_1 \mathrm{e}^{2\pi \mathrm{i} k_n a_1} + \cdots + m_r \mathrm{e}^{2\pi \mathrm{i} k_n a_r} \approx m_1 + \cdots + m_r = c_0.$$

For $l \in \mathbb{N}$, $l < r$, writing μ_i for $\lambda_i/\rho = \mathrm{e}^{2\pi \mathrm{i} a_i}$ as in the proof of the case when all a_i's are rational, this implies

$$c_l^{(n)} = m_1 \mathrm{e}^{2\pi \mathrm{i}(k_n + l) a_1} + \cdots + m_r \mathrm{e}^{2\pi \mathrm{i}(k_n + l) a_r} \approx m_1 \mu_1^l + \cdots + m_r \mu_r^l = c_l.$$

Moreover, this approximation becomes better with greater n, and in fact, Theorem 2 shows that $\lim_{n \to \infty} c_l^{(n)} = c_l$ for all l, $0 \leqslant l < r$. We know from the rational case that the determinant of the Vandermonde matrix with entries μ_i^l is nonzero which further implies that all c_l cannot be zero. Each c_l is the limit of $(c_l^{(n)})$, hence for the appropriate sequence of exponents $n = k_m + l$, $m \in \mathbb{N}$, the corresponding sequence $(c_l^{(n)})$ of coefficients at $\rho^n n^{\kappa - 1}$ has a nonzero limit. This finishes the proof of the divergence of $(p_j^* A^n q_k)$ in the last possible case.

The "if" part of the theorem is obviously true from Proposition 2, and so the proof of Theorem 1 is complete. $\square$

Acknowledgements

The author is greatly indebted to Professor Peter Šemrl for suggesting this area of research. He also wishes to express his gratitude to both referees for all the remarks that suggest him to simplify the last part of the proof.

References

[1] J.W.S. Cassels, An Introduction to Diophantine Approximation, Cambridge University Press, Cambridge, 1957.
[2] R.E. Hartwig, P. Šemrl, Constrained convergence, Rocky Mt. J. Math. 29 (1) (1999) 177–195.
[3] W.M. Schmidt, Diophantine Approximation, Springer, Berlin, 1980.
[4] W.M. Schmidt, Diophantine Approximations and Diophantine Equations, Springer, Berlin, 1991.
[5] F. Uhlig, Partial Convergence and Sign Convergence of Matrix Powers via Eigen c-ads, El. J. Linear Algebra 3 (1998) 75–92.